Research Article

Performance Comparison between the Specific and Baseline Prediction Models of Ecotoxicity for Pharmaceuticals: Is a Specific QSAR Model Inevitable?

Qingwei Bu ⓘ**, Qingshan Li** ⓘ**, Yun Liu** ⓘ**, and Chun Cai** ⓘ

School of Chemical & Environmental Engineering, China University of Mining & Technology-Beijing, Beijing 100083, China

Correspondence should be addressed to Qingwei Bu; qingwei.bu@cumtb.edu.cn

Assessing the ecotoxicity of pharmaceuticals is of urgent need due to the recognition of their possible adverse effects on nontarget organisms in the aquatic environment. The reality of ecotoxicity data scarcity promotes the development and application of quantitative structure activity relationship (QSAR) models. In the present study, we aimed to clarify whether a QSAR model of ecotoxicity specifically for pharmaceuticals is needed considering that pharmaceuticals are a class of chemicals with complex structures, multiple functional groups, and reactive properties. To this end, we conducted a performance comparison of two previously developed and validated QSAR models specifically for pharmaceuticals with the commonly used narcosis toxicity prediction model, i.e., Ecological Structure Activity Relationship (ECOSAR), using a subset of pharmaceuticals produced in China that had not been included in the training datasets of QSAR models under consideration. A variety of statistical measures demonstrated that the pharmaceutical specific model outperformed ECOSAR, indicating the necessity of developing a specific QSAR model of ecotoxicity for the active pharmaceutical contaminants. ECOSAR, which was generally used to predict the baseline or the minimum toxicity of a compound, generally underestimated the ecotoxicity of the analyzed pharmaceuticals. This could possibly be because some pharmaceuticals can react through specific modes of action. Nonetheless, it should be noted that 95% prediction intervals spread over approximately four orders of magnitude for both tested QSAR models specifically for pharmaceuticals.

1. Introduction

It is evident that humans benefit a lot from pharmaceuticals, a class of active chemicals that can interact with biological receptors in humans. However, also because of their reactive property [1], the dark side of the coin is that they could pose potential ecological threats to nontarget organisms after use and release into the environment [2, 3] because nontarget species like fishes could share similar biological receptors to drugs to humans [4–8]. Demonstrating and assessing the ecotoxicological relevance of the presence of pharmaceuticals as emerging contaminants in the environment is of increasing concern during recent years [2, 9–13].

To help understand the potential risk of pharmaceuticals to nontarget organisms, the first step is to determine their occurrence levels in various environmental matrices. Owing to the development of analytical technology for the detection of extremely low concentrations of pharmaceuticals, intensive monitoring has been conducted worldwide especially during recent decades [14–19]. While there is still a need for extensive monitoring practices considering the various temporal and spatial scales [14], it is a more urgent demand of a high amount of measured ecotoxicological data that are equivalently salient to assess the environmental risk of the identified and potential pharmaceutical contaminants.

Unfortunately, these kinds of ecotoxicity data for pharmaceuticals are highly scarce at present because the tests for assessing chemical ecotoxicity are usually costly, labor-intensive, and time-consuming. When the measured data are not available, Quantitative Structure Activity Relationship (QSAR) models could offer an effective alternative [20–22]. The QSAR model generally describes the

relationship between structural features and a toxicity endpoint, which is developed and validated by a rigorous procedure using the existing measured data [23, 24]. With a validated QSAR model, the toxicity of a chemical without measured data can be predicted. QSAR models have been extensively employed to estimate the hazard of pharmaceuticals, especially under the context of ranking and priority screening of highly risky candidates for further study and monitoring [25–30].

In the existing applications of QSAR models, the Ecological Structure Activity Relationships (ECOSAR) model was among the most commonly used one for estimating the ecotoxicity of pharmaceuticals [10]. ECOSAR is a well-established modelling tool based on the octanol-water partition coefficient (log Kow) by United States Environmental Protection Agency. This log Kow-based QSAR model worked well for predicting nonspecific narcosis toxicity that is a function of the tendency to dissolve into membranes. However, it could be a problem for pharmaceuticals. On the one hand, pharmaceuticals are complex chemicals with multiple functional groups, which will lead to uncertainty when designing an appropriate chemical class [21]. On the other hand, pharmaceuticals are always thought to cover a wide range of toxic mechanisms. Hence, specific modes of action (MOA) could be expected when they interact with nontarget organisms [10, 31], possibly making the narcotic toxicity prediction model inaccurate. Madden et al. [32] pointed out that toxicity predictions by ECOSAR for pharmaceuticals should be explained with caution, and the applicability domain (AD) should be evaluated carefully, considering the fact that ECOSAR was developed based on a dataset consisting of mainly industrial chemicals of simple structure with only a single functional group.

Recently, some authors argued that it is of necessity to develop the QSAR model of ecotoxicity specifically for pharmaceuticals as a substitute to ECOSAR, considering their special properties. As a result, some of QSAR models have been reported in literature [33–36]. In particular, Tugcu et al. [34] developed a QSAR model (hereafter called the "Tugcu's model") specifically for pharmaceuticals to estimate the acute toxicity to fish using multiple molecular descriptors. They concluded that their model could provide better predictions of toxicity than ECOSAR. In another study conducted by Sangion and Gramatica [33], QSAR models (hereafter called the "Sangion's model") of acute toxicity to three trophic levels of aquatic organisms were developed. The comparison of their model results with predictions obtained by using ECOSAR showed that ECOSAR was also outperformed. Nonetheless, their comparisons were limited to a small subset of pharmaceuticals, and more work should be done to further verify whether the better performance of QSAR models specifically for pharmaceuticals is solid.

The aim of the present study is further to clarify whether a specific QSAR model is indispensable for the accurate prediction of the ecotoxicity of pharmaceuticals. To this end, we predicted the ecotoxicity of a totally new set of pharmaceuticals for which the measured data are available, by using two externally validated QSAR models (i.e., Tugcu's model and Sangion's model) and ECOSAR. And then, the predictive ability of each pharmaceutical specific QSAR model was compared with that of ECOSAR, respectively, to demonstrate the goal of the present study. Possible factors affecting the performance of QSAR models were also discussed.

2. Methods

2.1. Description of QSAR Models Specifically for Pharmaceuticals. Two validated QSAR models, specifically for pharmaceuticals (i.e., Sangion's model and Tugcu's model), were selected for the comparison of predictive ability with ECOSAR. Sangion's model was integrated in a well-defined and user-friendly software package, QSARINS (v. 2.2.2) [37, 38]. It was developed for toxicity prediction of pharmaceuticals in three different endpoints, and species were selected: growth rate inhibition at 72 h (EC50) in Pseudokirchneriella subcapitata, immobilization at 48 h (EC50) in Daphnia magna, and mortality at 96 h (LC50) in *Pimephales promelas*. Tugcu's model was developed to predict the lowest 96-hour mortality (LC50) in various freshwater fish species. Details of the description for both models can be found in Sangion and Gramatica [33] and Tugcu et al. [34].

2.2. Collection of Measured Ecotoxicity Data. An initial list containing 663 compounds was compiled that consists of pharmaceuticals produced during 2011–2015 in China [39–43], by excluding vitamins, electrolyte solution additives, biochemical pharmaceuticals, as pharmaceuticals with a molecular weight over 2000 Da. The organometallics and inorganic compounds were also excluded in the initial list. Basic information was collected, including the Chemical Abstract Service Register Number (CASRN), molecular weight, molecular form, and therapeutic group. The CASRN and name of pharmaceuticals were cross validated by searching in multiple sources, including PubChem (https://pubchem.ncbi.nlm.nih.gov/) and SciFinder (https://scifinder.cas.org/), to assure the accurate citation of the CASRN. We searched measured ecotoxicity data in the ECOTOX database (https://cfpub.epa.gov/ecotox/) for the 663 pharmaceuticals in terms of CASRN, endpoints, and species following the purpose of this study. For the Tugcu's model, measured data of LC50 for various freshwater fish species were collected. In case of multiple measured toxicity values for the same specific endpoint, the lowest value was selected to maintain consistency with the data selection criteria used in model development.

To make an independent evaluation, pharmaceuticals that have measured toxicity data but had not been included in the training set of the tested QSAR models were selected. For comparison of predictive ability of Sangion's model with ECOSAR, a total of 41 toxicity data were identified, including 10 EC50s of Pseudokirchneriella

subcapitata, 27 EC50s of Daphnia magna, and 4 LC50s of Pimephales promelas. Regarding the case of Tugcu's model versus ECOSAR, we identified 43 LC50s of various fish species that can be used for comparison. All the data were compiled in Tables S1 and S2. The negative logarithm of the measured LC50 or EC50 values (pT) was calculated on a molar basis and used for comparison with the predicted ones.

2.3. Generation of Predicted Ecotoxicity Data.

Regarding Sangion's model, the structures of pharmaceuticals with measured data were sketched using Spartan 16 (v 2.0.7) and geometrically optimized by the semiempirical AM1 method. The predicted ecotoxicity data were generated by using QSARINS, in which the theoretical molecular descriptors were calculated by the PaDEL-Descriptor software (v 2.21). The leverage approach was used to check whether a given compound is within the scope of AD on the basis of the molecular descriptor space defined according to the training set during model development. The leverage value was automatically calculated by the QSARINS.

In the case of Tugcu's model, the structures of pharmaceuticals were also drawn using the Spartan 16 software. The conformation was optimized at the lowest energy by employing the semiempirical PM3 method. The molecular descriptors were calculated using the Dragon software (v 7.0.10), the same one as that by Tugcu et al. [34]. The predicted pTs were calculated using the established linear model with the reported molecular descriptors. The leverage approach was used to determine if a tested pharmaceutical is within the scope of AD of the developed model.

ECOSAR incorporated in EPI Suite (v 4.11) was run for the narcotic toxicity prediction by using CASRN or SMILES as inputs. As pharmaceuticals are not a traditional class of structurally or functionally similar compounds, ECOSAR would allocate some pharmaceuticals to more than one class, and consequently, there is more than one result for each of them. We took baseline toxicity predictions that handled pharmaceuticals as neutral organic, considering the goal of the present study. The tested pharmaceuticals were manually checked if they are inside the AD by checking log Kow values. The results showed that all the tested pharmaceuticals were in the AD of ECOSAR.

2.4. Measure of Predictive Ability.

To demonstrate whether the predictive ability of the tested QSAR models, a series of summary statistics were calculated for comparison: Pearson correlation coefficient, differences between predicted and measured pT (referred as minimum, mean, and maximum absolute residual, root of mean squared residual (RMSE), and calculated as equations (1)–(4)), and the percentage of pharmaceuticals with differences between predicted and measured pT was less than factors of 2, 5, 10, 100, and 1000. A linear regression analysis of predicted and measured pT was conducted, and the 95% prediction intervals were also calculated. The statistical analysis was accomplished in Microsoft Excel 2016.

$$\text{minimum absolute residual (log units)} = \min \left| pT_i - \widehat{pT}_i \right|, \tag{1}$$

$$\text{mean absolute residual (log units)} = \frac{1}{n} \sum_{i=1}^{n} \left| pT_i - \widehat{pT}_i \right|, \tag{2}$$

$$\text{maximum absolute residual (log units)} = \max \left| pT_i - \widehat{pT}_i \right|, \tag{3}$$

$$\text{RMSE} = \frac{1}{n} \sum_{i=1}^{n} \left(pT_i - \widehat{pT}_i \right)^2, \tag{4}$$

where pT_i and $\widehat{pT}_i$ are the measured and predicted values for negative logarithm of the LC50 or EC50.

3. Results

3.1. Sangion versus ECOSAR Model.

Statistics characterizing the predictive ability of Sangion's and ECOSAR models are summarized for comparison in Table 1. It is evident that the pharmaceutical specific model, Sangion's model, was more able to predict the toxicity trend than the generalized narcotic toxicity prediction model. The predictive ability of Sangion's model significantly surpassed ECOSAR no matter which parameter was used as a metric. Sangion's model, which is specific for pharmaceuticals, can give a prediction with uncertainty lower than two orders of magnitude for over 80% of the tested pharmaceuticals, outperforming that of ECOSAR (approximately 60%). If considering only pharmaceuticals in AD, this variable was further improved to around 95%.

For predictions produced by ECOSAR, high RMSE (2.7 log unit, Table 1) was observed in the present study, which is in accordance with [33]. However, our application of Sangion's model to a new set of pharmaceuticals showed an RMSE of 1.4 log unit (Table 1) that is significantly higher than that (0.6 log unit) applied to the prediction set during model development. This could be attributed to a few influence factors that will be discussed later.

The better performance of Sangion's model is also clearly demonstrated in Figure 1. Narrower 95% prediction intervals were found for Sangion's model, while the 95% prediction intervals were approximately eight orders of magnitude wide for ECOSAR. Furthermore, the regression analysis showed that Sangion's model in general underestimated the toxicity of the most toxic pharmaceuticals (Figure 1(a)), while the baseline toxicity model tended to underestimate the toxicity of most pharmaceuticals (Figure 1(b)).

3.2. Tugcu versus ECOSAR Model.

Similarly, the results showed that Tugcu's model outperformed ECOSAR, although a slightly higher Pearson correlation coefficient was obtained for ECOSAR than Tugcu's model (Table 2). Nonetheless, rest of the statistics clearly indicated the superiority of Tugcu's model to ECOSAR. In this application, only three pharmaceuticals that have measured toxicity data

TABLE 1: Summary statistics for performance comparison of Sangion's model versus ECOSAR.

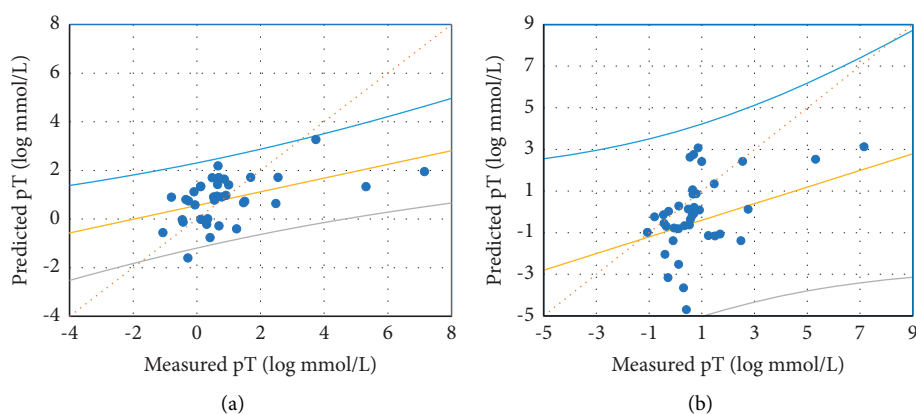| Measure of predictive ability | Sangion's model | | ECOSAR |
	Whole dataset	In AD	
Number of pharmaceuticals	41	36	41
Pearson correlation coefficient	0.422	0.483	0.272
Root mean square error, RMSE (log units)	1.630	1.422	2.744
Mean absolute residual (log units)	1.190	1.008	1.825
Minimum absolute residual (log units)	0.025	0.025	0.084
Maximum absolute residual (log units)	5.023	5.023	9.858
% pharmaceuticals <factor of 2	12.2	13.9	19.5
% pharmaceuticals <factor of 5	36.6	41.7	34.1
% pharmaceuticals <factor of 10	51.2	58.3	48.8
% pharmaceuticals <factor of 100	85.4	94.4	61.0
% pharmaceuticals <factor of 1000	92.7	94.4	85.4

(a)

(b)

FIGURE 1: QSAR model predictions versus measured toxicity. The best-fit linear regression and 95% prediction intervals are shown. (a) Sangion's model and (b) ECOSAR.

TABLE 2: Summary statistics for the performance comparison of Tugcu's model versus ECOSAR.

| Measure of predictive ability | Tugcu's model | | ECOSAR |
	Whole dataset	In AD	
Number of pharmaceuticals	43	40	43
Pearson correlation coefficient	0.345	0.386	0.442
Root mean square error, RMSE (log units)	1.600	1.608	2.486
Mean absolute residual (log units)	1.144	1.133	1.694
Minimum absolute residual (log units)	0.031	0.031	0.088
Maximum absolute residual (log units)	5.918	5.918	8.983
% pharmaceuticals <factor of 2	20.9	20.0	14.0
% pharmaceuticals <factor of 5	46.5	47.5	30.2
% pharmaceuticals <factor of 10	55.8	57.5	41.9
% pharmaceuticals <factor of 100	83.7	82.5	76.7
% pharmaceuticals <factor of 1000	95.3	95.0	88.4

were out of the AD of Tugcu's model. No obvious improvement in model performance was observed by restricting the dataset to include only pharmaceuticals in the AD.

Using the mean absolute residue (MAR) as a metric, Tugcu et al. [34] found in their application that their model and ECOSAR had MAR of 0.348 and 0.872, respectively, indicating that average difference between predicted and measured toxicity was less than one order of magnitude for both models. In contrast, in the present application, the

pharmaceutical specific model had a MAR of 1.14, indicating that the mean difference was slightly more than one order of magnitude (Table 2).

The better performance of Tugcu's model over ECOSAR was also demonstrated by the significantly narrower 95% prediction intervals as shown in Figure 2. Again, the 95% prediction intervals for predictions of ECOSAR were spread nearly eight orders of magnitude. It is also found from Figure 2 that Tugcu's model underestimated the toxicity for the most toxic compounds but overestimated the toxicity for
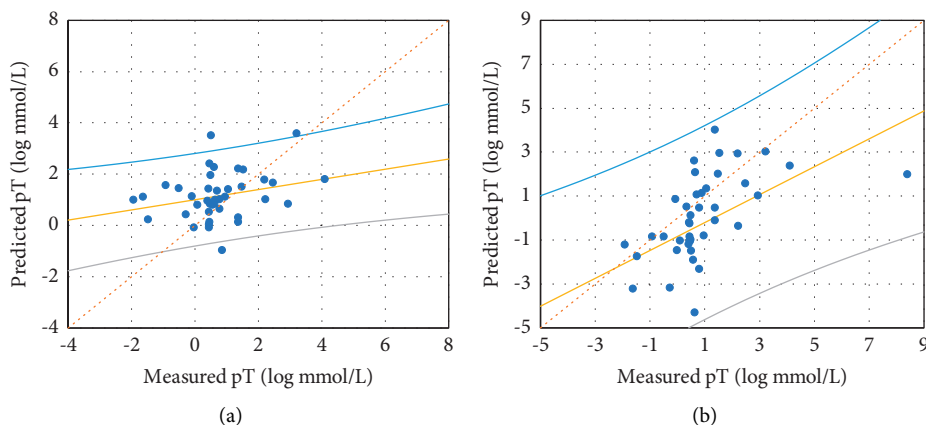
FIGURE 2: QSAR model predictions versus measured toxicity. The best-fit linear regression and 95% prediction intervals are shown. (a) Tugcu's model. (b) ECOSAR.

the least toxic compounds. And, the underestimation still existed for most pharmaceuticals when using ECOSAR to predict fish toxicity.

It should be noted that Tugcu's model was developed using ecotoxicity data of all fish species, which could influence the accuracy of model predictions. And, this was demonstrated to some extent when we did an intercomparison between the two pharmaceutical specific models. It is obvious that the predictive ability of Sangion's model was little better than that of Tugcu's model (Tables 1 and 2) although the intercomparison is not totally reasonable as they are applied to different pharmaceuticals.

4. Discussion

Our results clearly demonstrated that a better performance was observed for either Sangion's or Tugcu's model and the QSAR model specifically for pharmaceuticals, compared to the generalized baseline toxicity prediction model of ECOSAR. The present study further adds evidence that we need a QSAR model of ecotoxicity specifically for pharmaceuticals.

Not surprisingly, the model with more descriptors and sophisticated statistical methodology had a better overall performance. ECOSAR have been proved to be a valid tool for the prediction of ecotoxicity of the relatively simpler chemicals, but it could be problematic when being applied to predict the toxicity of chemicals with complex structures and functional groups as those of pharmaceuticals. In other words, the use of too simple equations cannot deal with the complexity of the phenomenon. Assigning a pharmaceutical, which is chemically complex, to one chemical class, means that only one functional subgroup is considered as responsible of the activity, and this can even be false. Modern QSAR models developed and trained on multiclass chemicals are by definition better able to work with more functional groups.

In the present study, the toxicity was generally underestimated by ECOSAR (Figures 1(b) and 2(b)). To further illustrate the predictive ability of ECOSAR, we also considered the lowest predictions in ECOSAR (worst case) in both comparisons. While the average difference between predicted and measured toxicity decreased by approximately 0.3 log unit, the ECOSAR still underestimated the toxicity for nearly half of the tested pharmaceuticals (Figures S1 and S2 in the Supplementary Materials).

The underestimation could be due to that the baseline toxicity, also known as the minimum toxicity, could not reflect the real ecotoxicity, and a specific MOA could exist for those reactive pharmaceuticals [44, 45]. Some authors pointed out that a specific MOA for pharmaceuticals could be indicated by the poor correlation between the acute toxicity and log K_{ow} [10, 31]. However, others have argued that log K_{ow} is a good predictor for most pharmaceuticals, indicating that they are the narcotic [46]. Sanderson and Thomsen [47] further concluded that a nonspecific MOA was suggested by the acute to chronic ratio (ACR) lower than 25 for nearly 70% of analyzed pharmaceuticals. However, it is complicated to recognize the true MOA of pharmaceuticals at present, especially for chronic MOA. The accuracy of using ACR to help determine the MOA of a compound heavily relied on the accurate determination of chronic toxicity. However, lacking high quality chronic data is indisputable, and therefore, high ACRs may also indicate a narcosis and conversely low ACRs may not necessarily indicate a nonspecific MOA [48]. Moreover, the standardized tests are designed for tests of mortality, which may overlook the sublethal effects even if there is any [46]. We have derived the threshold for two pharmaceuticals in different therapeutic groups and found that the traditional endpoints related to population mortality could not be the most sensitive one [49]. Adverse effects may occur via a possibly specific MOA that an organism would die from for a long run. It is found that, over two years, chronic exposure to 17α-ethylestradiol at extremely low concentrations (about 5 ng/L) led to a collapse of fathead minnow populations in a lake [44]. Further complicating this is that even for the same compound, MOAs may vary between organisms, at different life stages and at different exposure

concentrations [50], all of which would lead to uncertainties in ACRs' calculation.

Although the pharmaceutical specific QSAR models outperformed that of ECOASR, it should be noted that the 95% prediction intervals were approximately four orders of magnitude wide for both Sangion's and Tugcu's models (Figures 1 and 2).

Moore et al. [21] evaluated the performance of six QSAR packages that predict acute toxicity to fish for industrial chemicals and found even the best package which produced variation of four orders of magnitude, as our assessment for pharmaceuticals' toxicity prediction by the specific QSAR models. It is unclear if this is a systematic error that is caused by the existing uncertainty in measured toxicity data. We made a perfect assumption for the measured toxicity values that they are accurate, precise, and unbiased reflection of the true toxicity of pharmaceuticals to the aquatic organisms. Nonetheless, this is impossible considering that the measured toxicity data were determined under different testing protocols. Therefore, large residuals can be partly attributed to uncertainty that inherently exists in measured data.

On the other hand, it should also be kept in mind that no perfect model has ever been established. We examined the outlier substances that were defined as those substances that had differences between model predictions and measured values greater than 100 folds. Regarding Sangion's model, two outlier substances were identified: doramectin and niclosamide. They are both anticestodals that were not included in the training set of Sangion's model. Seven out forty-one substances were identified as outliers for Tugcu's model. The outlier substances include the antineoplastic drug, the estrogen, and the anticestodal drug. Again, they were rarely included in the training set during model development. And, estrogen is a well-proven contaminant that has specific MOA [51]. Therefore, limited coverage of pharmaceuticals of different therapeutic groups could be one of potential aspects influencing the performance of the pharmaceutical specific model.

It should be noted that the outlier substances were in the AD of the model, which means that the present molecular descriptors integrated in the model could not totally explain the toxicity variances for pharmaceuticals of different therapeutic uses. The specific QSAR models are "average" statistical results covering limited kinds of pharmaceuticals. They are still MOA-mixed QSAR models considering the fact that the pharmaceuticals of different therapeutic uses could possess different MOAs. It is hard to foresee the performance of a future developed model covering a wider range of pharmaceuticals belonging to various therapeutic groups. Either it will cover more predictors or will it be split to different models based on MOA to improve the model performance. The MOA-based QSAR has been advocated for a long time [52, 53]. However, there are great challenges to thoroughly understand the MOA of pharmaceutical's toxicity to nontarget organisms. As for pharmaceuticals, it has been argued that the mammalian pharmacology data could be potentially used to help understand their MOA [54]. The MOA-based QSAR model specifically for pharmaceuticals would be one of the approaches for an improved ecotoxicity prediction of these chemicals.

5. Conclusions

QSAR models have been recognized the most common tools for complementing the scarcity of toxicity data in scientific research and management. Our results showed that the generally used baseline toxicity prediction model, ECOSAR, usually underestimated the toxicity of pharmaceuticals in the present practice. The better performance of both Sangion's and Tugcu's models clearly demonstrated the necessity for QSAR models specifically for pharmaceuticals with complex structures and multiple functional groups. Modern QSAR models that are able to work with more functional groups are urgently needed. However, there is also a need for improving the performance of the present pharmaceutical-specific QSAR models either by covering a large pool of pharmaceuticals in different therapeutic groups or finally by developing an MOA-based model with the aid of existing pharmacology data.

Data Availability

The numerical data used to support the findings of this study are included within the article and the associated supporting information.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Supplementary Materials

Table S1: the measured (the lowest observed value) and predicted toxicity data for pharmaceuticals used for performance comparison of Sangion's model versus ECOSAR. Table S2: the measured (the lowest observed value) and predicted toxicity data for pharmaceuticals used for performance comparison of Tugcu's model versus ECOSAR. Figure S1: measured toxicity versus the lowest model prediction of ECOSAR for pharmaceuticals used for performance comparison of Sangion's model versus ECOSAR. The best-fit linear regression and 95% prediction intervals are shown. Figure S2: measured toxicity versus the lowest model prediction of ECOSAR for pharmaceuticals used for performance comparison of Tugcu's model versus ECOSAR. The best-fit linear regression and 95% prediction intervals are shown. (*Supplementary Materials*)

References

[1] S. A. Snyder, P. Westerhoff, Y. Yoon, and D. L. Sedlak, "Pharmaceuticals, personal care products, and endocrine disruptors in water: implications for the water industry," *Environmental Engineering Science*, vol. 20, no. 5, pp. 449–469, 2003.

[2] C. G. Daughton and T. A. Ternes, "Pharmaceuticals and personal care products in the environment: agents of subtle change?" *Environmental Health Perspectives*, vol. 107, no. Suppl 6, pp. 907–938, 1999.

[3] G. T. Ankley, B. W. Brooks, D. B. Huggett, a. J. P. Sumpter, and P. John, "Repeating history: pharmaceuticals in the environment," *Environmental Science and Technology*, vol. 41, no. 24, pp. 8211–8217, 2007.

[4] Ø. Andersen, V. G. H. Eijsink, and M. Thomassen, "Multiple variants of the peroxisome proliferator-activated receptor (PPAR) γ are expressed in the liver of Atlantic salmon (*Salmo salar*)," *Gene*, vol. 255, no. 2, pp. 411–418, 2000.

[5] J. G. Nickerson, S. G. Dugan, G. Drouin, and T. W. Moon, "A putative β2-adrenoceptor from the rainbow trout (Oncorhynuchus mykiss)," *European Journal of Biochemistry*, vol. 268, no. 24, pp. 6465–6472, 2001.

[6] F. Yamaguchi and S. Brenner, "Molecular cloning of 5-hydroxytryptamine (5-HT) type 1 receptor genes from the Japanese puffer fish, Fugu rubripes," *Gene*, vol. 191, no. 2, pp. 219–223, 1997.

[7] J. Fick, R. H. Lindberg, M. Tysklind, and D. G. J. Larsson, "Predicted critical environmental concentrations for 500 pharmaceuticals," *Regulatory Toxicology and Pharmacology*, vol. 58, no. 3, pp. 516–523, 2010.

[8] S. Owen, E. Giltrow, D. Huggett et al., "Comparative physiology, pharmacology and toxicology of β-blockers: mammals versus fish," *Aquatic Toxicology*, vol. 82, no. 3, pp. 145–162, 2007.

[9] B. I. Escher, R. Baumgartner, M. Koller, K. Treyer, J. Lienert, and C. S. McArdell, "Environmental toxicology and risk assessment of pharmaceuticals from hospital wastewater," *Water Research*, vol. 45, no. 1, pp. 75–92, 2011.

[10] K. Fent, A. Weston, and D. Caminada, "Ecotoxicology of human pharmaceuticals," *Aquatic Toxicology*, vol. 76, no. 2, pp. 122–159, 2006.

[11] C. G. Daughton, "Pharmaceutical ingredients in drinking water: overview of occurrence and significance of human exposure," in *Contaminants of Emerging Concern in the Environment: Ecological and Human Health Considerations*American Chemical Society, Washington, DC, USA, 2010.

[12] S. Comber, M. Gardner, P. Sörme, D. Leverett, and B. Ellor, "Active pharmaceutical ingredients entering the aquatic environment from wastewater treatment works: a cause for concern?" *The Science of the Total Environment*, vol. 613-614, pp. 538–547, 2018.

[13] J. Guo, K. Selby, and A. B. A. Boxall, "Assessment of the risks of mixtures of major use veterinary antibiotics in European surface waters," *Environmental Science and Technology*, vol. 50, no. 15, pp. 8282–8289, 2016.

[14] Q. Bu, B. Wang, J. Huang, S. Deng, and G. Yu, "Pharmaceuticals and personal care products in the aquatic environment in China: a review," *Journal of Hazardous Materials*, vol. 262, pp. 189–211, 2013.

[15] Q. Sui, X. Cao, S. Lu, W. Zhao, Z. Qiu, and G. Yu, "Occurrence, sources and fate of pharmaceuticals and personal care products in the groundwater: a review," *Emerging Contaminants*, vol. 1, no. 1, pp. 14–24, 2015.

[16] Y. Yang, Y. S. Ok, K.-H. Kim, E. E. Kwon, and Y. F. Tsang, "Occurrences and removal of pharmaceuticals and personal care products (PPCPs) in drinking water and water/sewage treatment plants: a review," *The Science of the Total Environment*, vol. 596-597, pp. 303–320, 2017.

[17] J. Wang and S. Wang, "Removal of pharmaceuticals and personal care products (PPCPs) from wastewater: a review," *Journal of Environmental Management*, vol. 182, pp. 620–640, 2016.

[18] M. Crane, C. Watts, and T. Boucard, "Chronic aquatic environmental risks from exposure to human pharmaceuticals," *The Science of the Total Environment*, vol. 367, no. 1, pp. 23–41, 2006.

[19] Q. Sui, W. Zhao, X. Cao et al., "Pharmaceuticals and personal care products in the leachates from a typical landfill reservoir of municipal solid waste in Shanghai, China: occurrence and removal by a full-scale membrane bioreactor," *Journal of Hazardous Materials*, vol. 323, pp. 99–108, 2017.

[20] M. Pavan and A. P. Worth, "Publicly-accessible QSAR software tools developed by the joint research centre," *SAR and QSAR in Environmental Research*, vol. 19, no. 7-8, pp. 785–799, 2008.

[21] D. R. J. Moore, R. L. Breton, and D. B. MacDonald, "A comparison of model performance for six quantitative structure-activity relationship packages that predict acute toxicity to fish," *Environmental Toxicology & Chemistry*, vol. 22, no. 8, pp. 1799–1809, 2003.

[22] M. T. D. Cronin, J. D. Walker, J. S. Jaworska, M. H. I. Comber, C. D. Watts, and A. P. Worth, "Use of QSARs in international decision-making frameworks to predict ecologic effects and environmental fate of chemical substances," *Environmental Health Perspectives*, vol. 111, no. 10, pp. 1376–1390, 2003.

[23] P. Gramatica, S. Cassani, P. P. Roy, S. Kovarich, C. Wei Yap, and E. Papa, "QSAR modeling is not push a button and find a correlation: a case study of toxicity of (Benzo-) triazoles on algae," *Molecular Informatics*, vol. 31, no. 11-12, pp. 817–835, 2012.

[24] OECD, *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models*, OECD Publishing, Paris, France, 2014.

[25] S. Ortiz de García, G. P. Pinto, P. A. García-Encina, and R. I. Mata, "Ranking of concern, based on environmental indexes, for pharmaceutical and personal care products: an application to the Spanish case," *Journal of Environmental Management*, vol. 129, pp. 384–397, 2013.

[26] H. Sanderson, D. J. Johnson, C. J. Wilson, R. A. Brain, and K. R. Solomon, "Probabilistic hazard assessment of environmentally occurring pharmaceuticals toxicity to fish, daphnids and algae by ECOSAR screening," *Toxicology Letters*, vol. 144, no. 3, pp. 383–395, 2003.

[27] K. Helwig, C. Hunter, M. McNaughtan, J. Roberts, and O. Pahl, "Ranking prescribed pharmaceuticals in terms of environmental risk: inclusion of hospital data and the importance of regular review," *Environmental Toxicology & Chemistry*, vol. 35, no. 4, pp. 1043–1050, 2016.

[28] K. Ji, E. J. Han, S. Back, J. Park, J. Ryu, and K. Choi, "Prioritizing human pharmaceuticals for ecological risks in the freshwater environment of Korea," *Environmental Toxicology & Chemistry*, vol. 35, no. 4, pp. 1028–1036, 2016.

[29] O. S. A. Al-Khazrajy and A. B. A. Boxall, "Risk-based prioritization of pharmaceuticals in the natural environment in

Iraq," *Environmental Science and Pollution Research*, vol. 23, no. 15, pp. 15712–15726, 2016.

[30] Z. Dong, D. B. Senn, R. E. Moran, and J. P. Shine, "Prioritizing environmental risk of prescription pharmaceuticals," *Regulatory Toxicology and Pharmacology*, vol. 65, no. 1, pp. 60–67, 2013.

[31] J. P. Seiler, "Pharmacodynamic activity of drugs and ecotoxicology—can the two be connected?" *Toxicology Letters*, vol. 131, no. 1, pp. 105–115, 2002.

[32] J. C. Madden, S. J. Enoch, M. Hewitt, and M. T. D. Cronin, "Pharmaceuticals in the environment: good practice in predicting acute ecotoxicological effects," *Toxicology Letters*, vol. 185, no. 2, pp. 85–101, 2009.

[33] A. Sangion and P. Gramatica, "Hazard of pharmaceuticals for aquatic environment: prioritization by structural approaches and prediction of ecotoxicity," *Environment International*, vol. 95, pp. 131–143, 2016.

[34] G. Tugcu, M. T. Saçan, M. Vračko, M. Novič, and N. Minovski, "QSTR modelling of the acute toxicity of pharmaceuticals to fish," *SAR and QSAR in Environmental Research*, vol. 23, no. 3-4, pp. 297–310, 2012.

[35] L. Jiang, Z. Lin, X. Hu, and D. Yin, "Toxicity prediction of antibiotics on luminescent bacteria, photobacterium phosphoreum, based on their quantitative structure-activity relationship models," *Bulletin of Environmental Contamination and Toxicology*, vol. 85, no. 6, pp. 550–555, 2010.

[36] K. P. Singh, S. Gupta, and N. Basant, "QSTR modeling for predicting aquatic toxicity of pharmacological active compounds in multiple test species for regulatory purpose," *Chemosphere*, vol. 120, pp. 680–689, 2015.

[37] P. Gramatica, S. Cassani, and N. Chirico, "QSARINS-chem: Insubria datasets and new QSAR/QSPR models for environmental pollutants in QSARINS," *Journal of Computational Chemistry*, vol. 35, no. 13, pp. 1036–1044, 2014.

[38] P. Gramatica, N. Chirico, E. Papa, S. Cassani, and S. Kovarich, "QSARINS: a new software for the development, analysis, and validation of QSAR MLR models," *Journal of Computational Chemistry*, vol. 34, no. 24, pp. 2121–2132, 2013.

[39] MIIT, *China Medical Statistics 2014*, Ministry of Industry and Information Technology of the People's Republic of China, Beijing, China, 2015.

[40] MIIT, *China Medical Statistics 2013*, Ministry of Industry and Information Technology of the People's Republic of China, Beijing, China, 2014.

[41] MIIT, *China Medical Statistics 2012*, Ministry of Industry and Information Technology of the People's Republic of China, Beijing, China, 2013.

[42] MIIT, *China Medical Statistics 2011*, Ministry of Industry and Information Technology of the People's Republic of China, Beijing, China, 2012.

[43] MIIT, *China Medical Statistics 2015*, Ministry of Industry and Information Technology of the People's Republic of China, Beijing, China, 2016.

[44] K. A. Kidd, P. J. Blanchfield, K. H. Mills et al., "Collapse of a fish population after exposure to a synthetic estrogen," *Proceedings of the National Academy of Sciences*, vol. 104, no. 21, pp. 8897–8901, 2007.

[45] R. Länge, T. H. Hutchinson, C. P. Croudace et al., "Effects of the synthetic estrogen 17α-ethinylestradiol on the life-cycle of the fathead minnow (Pimephales promelas)," *Environmental Toxicology & Chemistry*, vol. 20, no. 6, pp. 1216–1227, 2001.

[46] H. Sanderson and M. Thomsen, "Ecotoxicological quantitative structure-activity relationships for pharmaceuticals," *Bulletin of Environmental Contamination and Toxicology*, vol. 79, no. 3, pp. 331–335, 2007.

[47] H. Sanderson and M. Thomsen, "Comparative analysis of pharmaceuticals versus industrial chemicals acute aquatic toxicity classification according to the United Nations classification system for chemicals. Assessment of the (Q)SAR predictability of pharmaceuticals acute aquatic toxicity and their predominant acute toxic mode-of-action," *Toxicology Letters*, vol. 187, no. 2, pp. 84–93, 2009.

[48] J. Ahlers, C. Riedhammer, M. Vogliano, R.-U. Ebert, R. Kühne, and G. Schüürmann, "Acute to chronic ratios in aquatic toxicity-variation across trophic levels and relationship with chemical structure," *Environmental Toxicology & Chemistry*, vol. 25, no. 11, pp. 2937–2945, 2006.

[49] Q. Huang, Q. Bu, W. Zhong, K. Shi, Z. Cao, and G. Yu, "Derivation of aquatic predicted no-effect concentration (PNEC) for ibuprofen and sulfamethoxazole based on various toxicity endpoints and the associated risks," *Chemosphere*, vol. 193, pp. 223–229, 2018.

[50] M. Nendza and M. Müller, "Discriminating toxicant classes by mode of action: 2. Physico-chemical descriptors," *Quantitative Structure-Activity Relationships*, vol. 19, no. 6, pp. 581–598, 2000.

[51] A. C. Johnson and J. P. Sumpter, "Are we going about chemical risk assessment for the aquatic environment the wrong way?" *Environmental Toxicology & Chemistry*, vol. 35, no. 7, pp. 1609–1616, 2016.

[52] T. I. Netzeva, M. Pavan, and A. P. Worth, "Review of (quantitative) structure-activity relationships for acute aquatic toxicity," *QSAR & Combinatorial Science*, vol. 27, no. 1, pp. 77–90, 2008.

[53] A. P. Freidig and J. L. M. Hermens, "Narcosis and chemical reactivity QSARs for acute fish toxicity," *Quantitative Structure-Activity Relationships*, vol. 19, no. 6, pp. 547–553, 2000.

[54] M. J. Winter, S. F. Owen, R. Murray-Smith, G. H Panter, M. J Hetheridge, and L. B Kinter, "Using data from drug discovery and development to aid the aquatic environmental risk assessment of human pharmaceuticals: concepts, considerations, and challenges," *Integrated Environmental Assessment and Management*, vol. 6, no. 1, pp. 38–51, 2010.